

LISTING OF CLAIMS

A detailed listing of all claims that are, or were, in the present application, irrespective of whether the claim(s) remains under examination in the application are presented below. The claims are presented in ascending order and each includes one status identifier. Those claims not cancelled or withdrawn but amended by the current amendment utilize the following notations for amendment: 1. deleted matter is shown by strikethrough for six or more characters and double brackets for five or less characters; and 2. added matter is shown by underlining.

1. (Previously Presented) A method implemented on an information handling system for the automatic harvesting and qualification of dynamic database content comprising:

obtaining an initial categorization structure for organizing a plurality of subject areas of information;

obtaining a plurality of parametric information lists for optimizing operation to a user's requirements;

obtaining a candidate database listing having a plurality of candidate databases each having a collection of content;

acquiring a listing of a plurality of qualified databases from said candidate database listing by matching one of a candidate databases to said plurality of subject areas of information;

obtaining a query from a user, said query being associated with a subject area;

submitting said query to said plurality of qualified databases;

acquiring a collection of responsive content from said plurality of qualified databases;

indexing said responsive content to form an index of facilitating searching said collection of responsive content;

publishing a summary of said collection of responsive content for review by the user.

2. (Previously Presented) The method of claim 1, wherein said step of obtaining a plurality of parametric information lists further comprises:

obtaining an exclusion list providing a plurality of terms and sources to inhibit associations for said step of acquiring a collection of responsive content;

obtaining an inclusion list providing a plurality of terms and sources restricting associations for said step of acquiring a collection of responsive content;

obtaining a stop list providing a plurality of terms to be excluded for said step of indexing said responsive content.

3. (Previously Presented) The method of claim 1, wherein said step of acquiring a plurality of qualified databases further comprises:

- capturing an initial page from each one of said plurality of candidate databases;
- evaluating said initial page for relevancy to said each one of said subject areas of information;
- selecting databases according to relevance to said subject areas of information;
- associating said selected databases with said subject areas of information.

4. (Previously Presented) The method of claim 3, further comprising:

- obtaining a database relevancy parameter for restricting qualification of databases below a minimum threshold value;
- comparing the relevance of each initial page to said relevancy parameter;
- removing each candidate database with a relevancy below said minimum threshold value from qualification.

5. (Previously Presented) The method of claim 3, wherein said step of acquiring a plurality of qualified databases further comprises:

- submitting a query to each of said selected databases;
- capturing a plurality of pieces of responsive content provided by each of said selected databases;
- evaluating each of said plurality of pieces of responsive content for relevancy to said query;
- assigning a numerical score to each one of said plurality of pieces of responsive content, said numerical score representing a degree of relevance to said query;
- developing an aggregate score for each one of said select databases;

qualifying a portion of said selected databases based upon said aggregate score.

6. (Previously Presented) The method of claim 5, wherein said step of capturing a plurality of pieces of responsive content further comprises:

obtaining a content parameter limiting the number of pieces of content to be captured from each of said selected databases;

obtaining an initial weighting of each of said pieces of responsive content from said selected database;

selecting a quantity of pieces of responsive content limited by said content parameter such that pieces of responsive content with relatively greater initial weighting are selected before pieces of responsive content with a relatively lesser initial weighting.

7. (Previously Presented) The method of claim 1, wherein said step of acquiring a plurality of qualified databases further comprises:

capturing an initial page from each of said plurality of candidate databases;

evaluating said initial page for relevancy to said each of said subject areas;

obtaining a database relevancy parameter for restricting the qualification of databases below a minimum threshold value;

comparing the relevance of each initial page to said relevancy parameter;

removing each candidate database with a relevancy below said minimum threshold value from qualification;

selecting databases according to relevance to said subject areas;

submitting a query to said selected databases;

capturing a plurality of pieces of responsive content provided by said selected databases;

obtaining a content parameter limiting the number of pieces of content to be captured from said selected databases;

obtaining an initial weighting of each of said pieces of responsive content from said database;

selecting a quantity of pieces of responsive content limited by said content parameter such that pieces of responsive content with a relatively greater initial weighting are selected before pieces of responsive content with a relatively lesser initial weighting;

evaluating each of said plurality of pieces of responsive content for relevancy to said query;

assigning a numerical score to each of said plurality of pieces of responsive content, said numerical score representing a degree of relevance to said query;

developing an aggregate score for each of said selected databases;

qualifying apportion of said selected databases based upon said aggregate score to be polled for content;

associating said qualified databases with said subject areas.

8. (Previously Presented) The method of claim 1, wherein said step of acquiring a plurality of qualified databases further comprises:

analyzing an initial page from each one of said plurality of qualified databases for formatting;

determining an input location on said initial page for passing queries from said initial page to each one of said plurality of databases;

determining results locations for capturing search results returned from each one of said plurality of databases;

recording said input location and said results locations for use in formatting queries for each one of said databases.

9. (Original) The method of claim 1, wherein said step of acquiring a collection of responsive content further comprises:

comparing said piece of responsive content to each one of said subject areas in said initial categorization structure;

matching each piece of responsive content to subject areas based on relevance of the responsive content to the subject areas;

filtering matches to optimize said categorization structure.

10. (Original) The method of claim 9, wherein said step of filtering matches further comprises:

removing duplicate pieces of responsive content;

obtaining a population parameter for limiting a number of pieces of responsive content which may be matched to any one subject area;

obtaining an occurrence parameter for limiting a number of subject areas to which any one piece of responsive content may be matched;

restricting matches for each one of said subject areas according to said occurrence parameter and said population parameter.

11. (Original) The method of claim 9, wherein said step of filtering matches further comprises:

obtaining an exclusion list to inhibit matches based on predetermined words and sources;

obtaining an inclusion list to restrict matches based on predetermined words and sources;

matching each piece of responsive content with subject areas according to said exclusion list and said inclusion list.

12. (Original) The method of claim 9, further comprising:

creating a categorization file for recording matches between each piece of responsive content and each subject area;

saving said categorization file to a storage medium for use in searching said collection of responsive content.

13. (Original) The method of claim 1, wherein said step of indexing said responsive content further comprises:

obtaining a stop list providing a list of words not to be indexed;

parsing each piece of responsive content into constituent words;

eliminating words of said responsive content occurring on said stop lists;

recording a location of every occurrence of constituent words in said collection of responsive content.

14. (Original) The method of claim 1, wherein said step of publishing a summary further comprises:

determining if a summary is provided for each piece of said responsive content;

examining each piece of said responsive content for keywords associated with each subject area;

developing a keyword summary score for each piece of responsive content;

examining each piece of said responsive content for relevant extracts forming an extract summary;

- developing an extract score for each piece of responsive content;
- comparing said keyword summary score to said extract score for a summary composite score;
- selecting said keyword summary if a predetermined summary value is exceeded by said summary composite score;
- selecting said extract summary if a predetermined summary value if not exceeded by said summary composite score.

15. (Previously Presented) A system for the automatic harvesting and qualification of dynamic database content comprising:

- a computer system having a communication means for communicating with at least one other computer including a database to facilitate a two-way flow of information between said computer system and the at least one other computer;

- said computer system having a storage means for retention and recall of data communicated by or to the at least one other computer;

- said computer system having a processing means for executing multiple software modules and performing comparisons between a user supplied query and a plurality of documents found in at least one other computer;

- an index for storing a plurality of pre-approved internet sites to be included in a series of queries;

- a configuration module adapted for translating a generic query into site-specific dialects such that a single user defined query may be directed to multiple sites automatically;

- a selection module adapted for characterizing said plurality of documents returned by the database of the at least one other computer associated with said user defined query;



a results index to allow for rapid recovery of specific portions of any one of said plurality of documents characterized by said selection module; and

a generator module for automatically generating at least one results page for the user conveying information associated with any one of said plurality of documents associated with said query.

16. (Previously Presented) A method implemented on an information handling system for the automatic harvesting and qualification of dynamic database content comprising:

obtaining an initial categorization structure for organizing a plurality of subject areas of information;

obtaining a plurality of parametric information lists for optimizing operation to a user's requirements;

obtaining a candidate database listing having a plurality of databases each having a collection of content, each one of said plurality of databases providing a dynamic response based upon a specific query;

acquiring a listing of a plurality of qualified databases from said candidate database listing by matching each one of a candidate databases to said plurality of subject areas;

obtaining a query from the user, said query being associated with a subject area;

submitting said query to said plurality of qualified databases;

acquiring a collection of responsive content from said plurality of qualified databases;

indexing said responsive content to form an index of facilitating searching said collection of responsive content;

publishing a summary of said collection of responsive content for review by the user.

17. (Previously Presented) The method of claim 16, wherein said step of acquiring a plurality of qualified databases further comprises:

- analyzing an initial page from each one of said plurality of qualified databases for formatting;

- determining an input location on said initial page for passing queries by said initial page to each one of said plurality of databases;

- analyzing a results page from each one of said plurality of qualified databases for formatting;

- determining results locations on said results page for capturing search results returned from each one of said plurality of databases;

- recording said input location and said results locations for use in formatting queries for each one of said databases.

18. (Previously Presented) The method of claim 1, wherein said step of acquiring a plurality of qualified databases further comprises:

- analyzing an initial page from each one of said plurality of qualified database for formatting;

- determining an input location on said initial page for passing queries by said initial page to each one of said plurality of databases;

- analyzing a results page from each one of said plurality of qualified databases for formatting;

- determining results locations on said results page for capturing search results returned from each one of said plurality of databases;

- recording said input location and said results locations for use in formatting queries for each one of said databases.

19. (Previously Presented) The method of claim 1, wherein said step of acquiring a listing of a plurality of qualified databases further comprises acquiring a listing of a plurality of qualified databases each generating dynamic responses based upon a user query.

20. (Previously Presented) The method of claim 1, further comprising the step of generating at least one summary comprising at least one extract of relevant content taken directly from an associated at least one item in said collection of responsive content from said plurality of qualified databases.

21. (Previously Presented) The method of claim 15, wherein characterizing said plurality of documents includes categorizing said plurality of documents.

22. (Previously Presented) The method of claim 15, wherein characterizing said plurality of documents includes evaluating said plurality of documents using parametric information lists.

23. (Previously Presented) The method of claim 15, wherein the selection module is further adapted to use one or more of the following to characterize said plurality of documents:

an exclusion list to provide a plurality of terms and sources to inhibit associations for acquiring a collection of responsive content;

an inclusion list to provide a plurality of terms and sources restricting associations for said step of acquiring a collection of responsive content; and

a stop list to provide a plurality of terms to be excluded.

24. (Previously Presented) The method of claim 15, wherein the processing means comprises:

a first evaluation means to compare a listing of candidate databases to a categorization structure to determine a listing of qualified databases;

a configuration means to create a configuration file to serve as a translator between a generic query and each qualified database;

a responsive determination means to determine responsive databases among the qualified databases;

an evaluation means to evaluate the responsive pages for relevance; and

an acquisition means to acquire documents from the responsive pages.

25. (Previously Presented) The method of claim 24, wherein the acquisition means is adapted to:

submit one or more queries to the responsive databases;

determine whether documents returned by the queries are stored on a central database;

and

capture and store documents that do not exist in the central database in the central database.